Conversational Voice Agents are Preferred and Lead to Better Driving Performance in Conditionally Automated Vehicles

Manhua Wang
Virginia Tech, Blacksburg, Virginia, United States
manhua@vt.edu

Seul Chan Lee
Gyeongsang National University, Jinju, Republic of Korea
seulchan@gnu.ac.kr

Genevieve Montavon
Virginia Tech, Blacksburg, Virginia, United States
genmontavon@vt.edu

Jiakang Qin
Virginia Tech, Blacksburg, Virginia, United States
qjk8888@vt.edu

Myounghoon Jeon
Virginia Tech, Blacksburg, Virginia, United States
myounghoonjeon@vt.edu

ABSTRACT

In-vehicle intelligent agents (IVIAs) can provide versatile information on vehicle status and road events and further promote user perceptions such as trust. However, IVIAs need to be constructed carefully to reduce distraction and prevent unintended consequences like overreliance, especially when driver intervention is still required in conditional automation. To investigate the effects of speech style (informative vs. conversational) and embodiment (voice-only vs. robot) of IVIAs on driver perception and performance in conditionally automated vehicles, we recruited 24 young drivers to experience four driving scenarios in a simulator. Results indicated that although robot agents received higher system response accuracy and trust scores, they were not preferred due to great visual distraction. Conversational agents were generally favored and led to better takeover quality in terms of lower speed and smaller standard deviation of lane position. Our findings provide a valuable perspective on balancing user preference and subsequent user performance when designing IVIAs.

CCS CONCEPTS

• **Human-centered computing**; • **Human computer interaction (HCI)**; • **Empirical studies in HCI**;

KEYWORDS

in-vehicle intelligent agent, conditionally automated driving, takeover performance, situation awareness

ACM Reference Format:
Manhua Wang, Seul Chan Lee, Genevieve Montavon, Jiakang Qin, and Myounghoon Jeon. 2022. Conversational Voice Agents are Preferred and Lead to Better Driving Performance in Conditionally Automated Vehicles. In *14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22), September 17–20, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3543174.3546830

AutomotiveUI '22, September 17–20, 2022, Seoul, Republic of Korea

ACM ISBN 978-1-4503-9415-4/22/09.
https://doi.org/10.1145/3543174.3546830

1 INTRODUCTION

Intelligent agents (IAs), who autonomously and intelligently help users with given tasks [39], have been widely used as voice assistants in people's daily lives. Driving is one of the promising use contexts that can greatly benefit from IAs. With IAs, drivers experience reduced visual distraction when interacting with in-vehicle infotainment systems [1, 33]. However, driving involves more complicated and safety-critical tasks; thus, solely query-and-answer or request-and-execute interaction mode in conventional IAs may no longer provide effective assistance. In-vehicle intelligent agents (IVIAs) are expected to proactively engage in driving tasks on a wider spectrum to secure road safety, supporting drivers in both driving-related and non-driving-related activities [51].

As vehicle automation gradually progresses to maturity, drivers' responsibilities will eventually fade away, switching them from active operators to passive monitors. At that time, the roles of IVIAs will transfer from ensuring safety to promoting drivers' in-cabin experience and their understanding of automation systems. IVIAs can provide vehicle- and scenario-related information. This information is critical to constructing explainable automation systems, helping drivers establish appropriate mental models, calibrate their trust towards automated vehicles (AVs), and facilitate technology acceptance. In addition to semantic contents, social attributes of IVIAs—such as speech style and embodiment—can also be manipulated to facilitate trust in advanced AVs [5, 22, 25]. Agents sharing more social attributes (e.g., conversational speech style, humanoid appearance) are more anthropomorphic [14, 22, 32] and can elicit a better driver experience. Conversational speech style is typically favored in autonomous vehicles (i.e., full AVs), with higher ratings on positive user perceptions such as trust [24, 25, 27], warmth [52], perceived usefulness, and ease of use [30], and lower ratings on negative perceptions such as cognitive demand and annoyance when interacting with IVIAs [52]. On the contrary, research on the influences of the embodiment has not reached an agreement. Embodied agents were perceived with higher competence and were preferred as companions in full AVs [32, 52], but they were not always preferred even without introducing extra visual distraction [8]. In any case, findings from existing empirical studies have shed light on the promising future of IVIAs in fully autonomous driving contexts.

However, full AVs are yet to enter the market soon, which will leave conditional AVs on the road for the next decades [26]. IVIAs in conditional AVs are still expected to support driving-related tasks as a functional objective to release drivers' burden [31] because drivers are required to take over the control from the automation systems and intervene in critical events when AVs reach their operational limits. This automation-to-driver authority transition process is called a takeover process [36]. The smoothness and safety of the takeover process are primarily determined by the timeliness of the takeover requests (TORs) and the effectiveness of the driver intervention [36]. If designed properly, IVIAs can play active roles in providing timely TORs and assisting drivers in negotiating critical events.

Existing research has investigated the design components of IVIAs that contribute to the timeliness of TORs. Simply adding IVIAs increased the likelihood of drivers making timely reactions [34]. The time-to-collision at the time of the TOR, also known as the lead time, also significantly impacts the driver reaction time as well as the lateral and longitudinal posttakeover driver intervention [36]. Generally, longer lead times are associated with longer takeover times [36, 58], while the crash rate shows a U-shaped relationship as the lead time increases [58]. In addition to the temporal variable, varying signal words, the tone, or the loudness of speech can convey different levels of perceived urgency, which further influences takeover reaction times [43, 45, 54, 58].

IVIAs can also assist drivers with intervention strategies because they are capable of and versatile in carrying information. Effective driver interventions are guarded by appropriate situation awareness (SA) in drivers, which is largely compromised in conditional AVs. Drivers in conditional AVs—as passive monitors—have declined alertness caused by task disengagement, low workload, or passive fatigue [34, 50]. Vehicle- and scenario-related information provided by IVIAs can keep drivers in the loop, improve their SA, and prompt future actions [16, 34, 35, 38]. For instance, the *"how"* message announces the vehicle's current action, and the "*why*" message explains the reason for vehicle decisions [21], while the "*what will*" message provides further recommendations in reaction to the scenario [9]. The "*what will*" message was perceived as more useful and easier to use compared to others in conditional AVs [9]. These semantic contents are important to improve drivers' SA and are critical to communicating automation capabilities and limitations, helping calibrate trust and avoid takeover failure due to system misuse [28, 41]. Social attributes of IVIAs are also beneficial in trust calibration in conditional AVs, which can further impact the effectiveness of driver intervention.

Exactly due to their versatility, IVIAs should be carefully designed. Subtle differences in speech interactions can significantly impact driver preference [53], and most importantly, can influence trust and reliance [28, 56]. Additionally, IVIAs can elicit distractions and annoyance due to their high social presence. While the influences of semantic contents on the takeover process have been researched widely in conditional AV contexts, the examination of social attributes of IVIAs has been primarily practiced under full AV conditions. Considering that drivers' responsibilities alter between full and conditional AVs [47], findings under fully autonomous driving conditions cannot be widely generalized in conditionally automated driving contexts. Drivers in full AVs are not responsible for any driving tasks; instead, they are open to other non-driving-related interactions. However, drivers in conditional AVs are expected to make timely and effective actions when asked. Thus, distractions and impairment in conditional AVs can be detrimental. In this case, how the social attributes of IVIAs influence driver perception and further impact the takeover process in conditional AVs remains ambiguous. To address this issue, the present study aims to systematically evaluate the effects of speech style and embodiment—as two characteristics contributing to social attributes of IVIAs— and their interaction effect on driver perception and takeover performance in conditional AVs. Two representative speech styles—informative vs. conversational—were used to create divergent perceptions: the informative style sounds commanding due to its simplicity and directness, while the conversational one is more suggestive and can create a feeling of being cared for [29]. The influence of the absence or presence of a physical body (voice-only vs. robot) were examined to understand possible distraction introduced by embodied agents, while also maximizing the anthropomorphism provided by a humanoid robot [46, 59]. We adopted a within-subjects factorial design, attempting to detach the effects of two social attributes from each other and provide an unambiguous view of their influential mechanisms. Based on the existing evidence, we further hypothesize that:

H1: Drivers will prefer voice-only agents without visual distraction in conditional AVs over robot agents, while conversational speech style will also be preferred in this context to enhance driver experience and engagement.

H2: Drivers accompanied by conversational agents will have better performance, specifically:

H2.1: Drivers will have better SA.

H2.2: Drivers will demonstrate safer takeover performance.

This research study leads to the following unique contributions. First, shifting in driver experience, especially their preference towards IVIAs, when compared with IVIA design in full AVs, uncovers and supports the dynamic user needs and requirements as the levels of automation alter. Further, user perception of the driver-agent interaction provides insights on how to design agents to promote driver experiences in conditional AVs, laying the foundation for comparing IVIA designs and preferences across different levels of automation. Finally, findings on the influence of different types of IVIAs on the subsequent takeover process can help with a balanced design between subjective preferences and unintended consequences.

2 METHODS

2.1 Participants

Twenty-four participants (7 females) aged between 19 and 33 years old (*Mean* = 23.12, *SD* = 4.49) with normal or corrected-to-normal vision participated in our study. All participants had valid driver's licenses and had an average driving experience of 5.93 years (*SD* = 3.37), with an average driving frequency of 4.67 days per week (*SD* = 2.22). Two participants had experience in partial AVs (i.e., Tesla full self-driving) before their participation.

2.2 Experimental apparatus and stimuli

We conducted a driving simulator study in a motion-based driving simulator (Nervtech$^{\text{TM}}$, Ljubljana, Slovenia), which consisted of three 48" displays that created a 120° horizontal field of view, an adjustable seat, a steering wheel, pedals for gas and brake, and

Figure 1: Experimental setup with the NervTech driving simulator and Nao.

surrounding sound equipment. Driving scenarios programmed in SCANeR studio were designed to simulate an SAE Level 3 Conditional Driving Automation [47]. We developed four driving scenarios, including straight and curved roads with traffic, traffic signals and signs, and other road users (e.g., pedestrians and other vehicles). The simulated ego vehicle had longitudinal and lateral control while navigating along a predefined route and handling limited road events such as stopping at a red light and crossing a controlled intersection. When the system reached its limitation (e.g., limited visibility due to weather, surprising event), a speech takeover request (TOR) along with a visual notification on the navigation panel would be issued to mandate the participant to take over control of the vehicle. The participant deactivated autonomous driving by either using a toggle attached to the steering wheel or pressing the brake. Upon exiting the takeover event zone, the system prompted the participant to reengage the automated mode (i.e., "Please reengage the auto-drive"). Each scenario consisted of four takeover events and four non-takeover events, lasting approximately seven minutes. The route and order of events were different among four scenarios to minimize learning effects.

Speech messages regarding road events and TORs used in this study were converted via the text-to-speech engine in Amazon Polly (name: Joanna, gender: female, nationality: USA). A humanoid robot, NAO (V6 standard edition, height: 22", width: 10.8"), was used under the embodied agent conditions. Figure 1 presents the experimental setup for the driving simulator and NAO. To capture participants' gaze fixation during the study, an eye-tracking device—Tobii Pro Glasses 2—with a sampling rate of 50 Hz was used.

2.3 Experimental design

This study adopted a 2 (Speech style: informative vs. conversational) x 2 (Embodiment: voice-only vs. robot) within-subjects factorial design. Thus, four types of in-vehicle intelligent agents (IVIAs) were evaluated in this study: informative voice agent (IVA), informative robot agent (IRA), conversational voice agent (CVA), and conversational robot agent (CRA). Each participant was accompanied by all four agents in four different driving scenarios, respectively. The order of the agent conditions was counterbalanced across participants and with the matching scenarios. Thus, the same agent was not always used in the same scenario to avoid the scenario as a confounding variable.

The IVIAs issued TORs and provided information regarding road events (see Table 1 for a complete list of events and scripts). All road events shared similar elements across four scenarios but were placed at different locations along the travel route to avoid learning effect. Thus, the difficulty of takeover events remained similar. While informative agents present information in a simple manner without additional information other than road events (e.g., "Exit ahead"), conversational agents communicate the message in a dialogue style (e.g., "We are entering a new road."). The TORs remained the same between informative and conversational styles to control the message length as a confounding variable that could impact information processing time and further influence the takeover reaction time under emergency situations. The length of takeover request ranged from 2.41 to 3.00 seconds (*Mean* = 2.64, *SD* = 0.27) across four events. For other road events, the length of informative messages ranged from 0.86 to 1.31 seconds (*Mean* = 1.09, *SD* = 0.21) and the length of conversational messages was between 1.13 to 4.08 seconds (*Mean* = 1.97, *SD* = 1.39).

The lead time was 4.5 seconds in this study. We selected a relatively limited time budget for the following reasons. First, our pilot study with a 7-second lead time for takeover events indicated a lower level of task difficulty, which led to performance degradation or passive fatigue due to boredom. Thus, we increased the task difficulty by limiting the time budget to keep participants' active engagement. Second, an empirical study showed a 4.5-second time budget led to a minimum crash rate and brake-to-maximum reaction time to speech warnings among the lead time shorter than 7 seconds [58].

2.4 Dependent measures and analysis

Both subjective and objective dependent measures were considered. Subjective measures included questionnaires to evaluate driver-agent interaction experience and driver preference. Objective measures included situation awareness, eye-tracking measures, and takeover performance. The following sections introduced the dependent measures separately and explained the analysis method afterward.

2.4.1 *Subjective measures.*

Subjective measures included ratings from three questionnaires collecting driver perception on the accompanying agent: the modified Subjective Assessment of Speech System Interfaces (SASSI) [17] (the Habitability and Speed subscales were removed due to their irrelevance to our agent setting), and the Scale of Trust in Automated Systems [20]. In addition, participants' preferences and reasons behind their first and least preferred agents were also asked. A two-way repeated-measures analysis of variance (ANOVA) was conducted to understand the influence of speech style and embodiment and their interaction effect on each factor of driver-agent interaction questionnaires. A Chi-square test for each preference rank was conducted to identify differences in preferred agents.

2.4.2 *Situation awareness.*

Drivers' situation awareness (SA) was evaluated using the Situation Awareness Global Assessment Technique (SAGAT) [13]. To develop the SA queries, we conducted a Goal-Directed Task Analysis (GDTA) to identify the SA requirements [12] needed for drivers under conditional automation to make decisions if a TOR was issued. With a list of SA requirements, six queries were constructed for each freeze point in the driving scenario, consisting of two queries for each level of SA: perception,

Table 1: Driving events and scripts in scenario

Event List	Informative Script List	Conversational Script List
Road construction	Take over immediately. Road construction ahead.	
Fog	Take over immediately. Fog ahead.	
Jaywalking	Take over immediately. Jaywalker ahead.	
Tunnel	Take over immediately. Tunnel ahead.	
Exit or enter a new road	Exit ahead.	We are entering a new road.
Waiting for a traffic signal	Red light ahead.	We are waiting for the signal to turn green.
Turning left/right	Turning left/right ahead.	We are turning left/right.
Two-way stop intersection	This is a two-way stop.	We've reached a two-way stop. We are waiting for other cars to go first.

comprehension, and projection [10]. Each query had four options with only one correct option. Participants' overall accuracy rate for each query was calculated. Queries with a lower than 25% accuracy rate ($N = 2$) of guess level were removed from further analysis. Then, the frequency of correctness (% correct) was calculated for each scenario. Because data from queries scored as correct or incorrect were binomial, the arcsine-root-square transformation was applied as a correction factor to allow the ANOVA tests [4, 11].

2.4.3 Eye-tracking measures. The eye-tracking data were collected and stored in the eye-tracking device. We primarily focused on gaze fixation to identify any potential distraction due to introducing an embodied agent. Specifically, we calculated distraction fixation frequency and total distraction duration.

Gaze fixations on predefined areas of interest (AOIs) were identified in the Tobii Pro Lab software (v1.152) and classified into two primary categories: driving-related fixations and distraction fixations. Driving-related fixations included fixations to the road, other road users, road signs and signals, rear-view mirrors, and the instrument panel. Distraction fixations included fixations to NAO in robot agent conditions and personal devices. The distraction fixation frequency and total distraction duration was calculated using the formula below [52, 57]:

$$Distration\ Fixation\ Frequency = \frac{FixationCount_{distraction}}{FixationCount_{total}} \quad (1)$$

The total distraction duration was the sum of all distraction fixation duration in seconds. After the calculation, a two-way repeated-measures ANOVA was performed to identify the effect of speech style and embodiment.

2.4.4 Takeover performance. Takeover performance was further divided into takeover time and quality (Table 2). Takeover time was defined as the time interval between the issue of TOR and the automation deactivation [7, 49], either by using the toggle or pressing the brake. Takeover quality metrics examined in the present study were speed-related measures (maximum, minimum, and average speed), maximum lateral acceleration, and standard deviation of lane position (SDLP; excluding the construction event, which required lance changing). Smaller values in speed, acceleration, or SDLP indicate smoother and safer takeover reactions. All takeover quality metrics were calculated during the manual control period between the time participants initiated the manual control and the time when they exited the takeover zone, which was a fixed point marked in the scenario and did not depend on participants' maneuver variation.

Each participant experienced four takeover events in each of four scenarios, resulting in a total of 384 data points for each takeover performance measure. Values exceeding six standard deviations for each measure were revisited and corrected if a programming error was detected or removed if a true outlier was determined. No more than 3% of the total number of data points were excluded from each measure, with the maximum lateral acceleration having the largest number of points removed *($n = 11$)* – mainly because of the simulation running error. Because the construction takeover event required a lane-changing maneuver, measures for this event were analyzed separately. The measures for the other three events were integrated and analyzed together. A two-way repeated ANOVA was conducted to determine the effect of speech style and embodiment on each measure for each event category.

2.5 Procedure

Upon arrival at the lab space, participants signed the consent form for this study approved by the university's Institutional Review Board. Participants were explained that there were four driving scenarios with conditional automation where they would not operate the vehicle for most of the time. However, if the system asked them to do so, they had to take over the control and drive for some time before handing over the control back to the vehicle when prompted. To simulate a natural driving situation, participants were allowed to do any tasks of their choice during the drives, but they must be ready to take over the control when asked. Participants were also informed of the presence of the IVIAs, who would issue TORs and provide information regarding other road events. Before the formal drives, a simulation sickness test was administrated [15], where a self-comfort checklist was completed before and after the 5-minute test drive. During this process, participants experienced sample takeover events and a pause for SAGAT with sample queries, while also familiarizing themselves with the system control and simulated scenarios. Then, demographical information was collected if they were not suspected of simulation sickness. Before the first drive, the eye-tracking glasses were put on participants and appropriately calibrated. During the drive, the experimenter paused the scenario at two certain points—differed for each driving scenario—and administrated the SAGAT queries. After finishing each drive,

Table 2: Takeover performance measures

Category	Dependent Measures	Unit	Definition
Temporal measures	Takeover time	Seconds	Time between TOR and automation deactivation.
Takeover quality*	Max/Min/Average speed	m/s	Maximum, minimum, or average speed during the takeover event after automation deactivation.
Takeover quality	Maximum lateral acceleration	m/s2	Maximum lateral acceleration during the takeover event after automation deactivation.
Takeover quality	SDLP	Meters	Standard deviation of the lateral distance of the ego vehicle regarding the middle of the lane.

* All takeover quality was calculated within the manual driving time frame within the takeover event zone defined along the route.

Table 3: Subjective ratings on driver-agent interaction

Scale	Items/Factors	Agent Type				Main Effect	
		IVA	CVA	IRA	CRA	Speech Style	Embodiment
SASSI	System Response Accuracy	4.96	4.82	5.28	5.17	-	V < R *
	Likability	4.85	4.79	4.74	5.07	-	-
	Cognitive Demand	3.18	3.03	2.98	3.13	-	-
	Annoyance	3.33	3.10	3.43	3.16	-	-
Trust in Automation	Trust	5.14	5.00	5.36	5.21	-	V < R *

* p < .05, I = informative, C = conversational, V = voice-only, R = robot.

participants completed the subjective questionnaires that collected their driver-agent interaction experience. After completing all conditions, participants ranked their preference towards four types of IVIAs and their reasons for the first and least preferred agents. The experiment lasted approximately 90 minutes.

3 RESULT

3.1 Driver-Agent Interaction

Table 3 summarizes the average rating score for each questionnaire under each condition. Embodiment showed a significant main effect in System Response Accuracy scale: $F(1, 23) = 5.51$, $p < .05$, $\eta_p^2 = .19$, and in Trust in Automation scale: $F(1, 23) = 4.35$, $p < .05$, $\eta_p^2 = .16$. Robot agents were perceived to have higher system response accuracy and trust than voice-only agents, regardless of their speech style.

No interaction effect between speech style and embodiment was identified in any factors.

3.2 Agent preference

Table 4 presents the preference ranking and distribution for each type of agent. A significant difference in participants' 1st preferred agent was found: $\chi^2(3) = 8.33$, $p < .05$. The conversational voice agent was preferred the most. When participants explained the reasons behind their first preferred agent, explanatory or instructive (N = 7), less distracting (N = 5), friendly (N = 4), and human-like (N = 3) were the most frequently mentioned perceived characteristics of the conversational voice agents:

> "The conversational voice agent provides the right information needed without being **distracting** as it may be with the robot or **condescending** with the strictly informative voice agent." (P5)

Table 4: Preference ranking for all agent conditions

Preference	Agent Type			
	IVA	CVA	IRA	CRA
1st	6	11	1	6
2nd	6	6	6	6
3rd	5	5	8	6
4th	7	2	9	6
Average Score	2.54	1.89	3.04	2.50
SD	1.18	1.02	0.91	1.14

Note: unit – number of participants

> "... remain **relaxed and focused** when having a voice that spoke in a **friendly** manner." (P10)

> "Being conversational engages the driver and **provides correct instructions** to prevent any human errors." (P12)

> "It is better to drive when the voice is more **human-like** and less of a robot. Also, the conversation makes it more **lifelike** as well. And **explains more** of what you should do." (P16)

When participants were asked to reason their least preferred agents, impolite/commanding/robotic (N = 7) and lack of information (N = 6) were raised frequently for informative agents, while distracting (N = 4) and uncomfortable (N = 3) were frequently mentioned for robot agents:

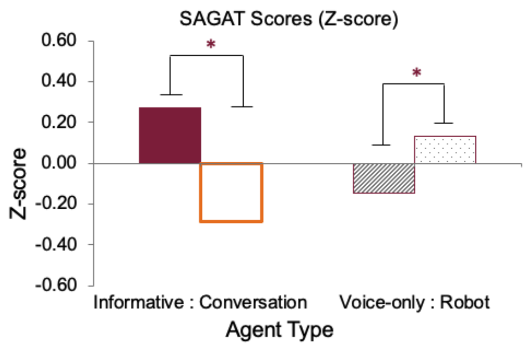

Figure 2: Standardized SAGAT scores across conditions (* $p <$ .05)

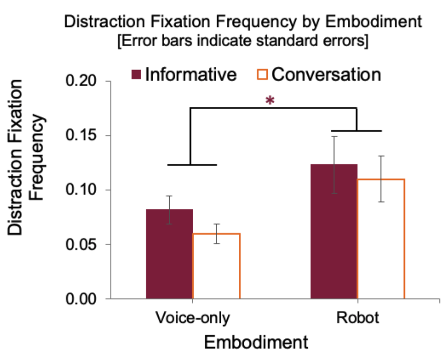

Figure 3: Distraction fixation frequency among all conditions (* $p < .05$) [Error bars represent standard errors]

> "The voice was very **robotic** and **unpleasant** to listen to. Also, the robot sometimes stares at me, which was **distracting**." (P9, 4th choice: IRA)
>
> ". . . the informative option was less ideal because it simply stated things that were obvious from the surroundings already." (P12, 4th choice: IRA)
>
> "I feel like the informative voice agent was telling me what to do and **not in a polite way**. It felt like a backseat driver, which can get annoying." (P20, 4th choice: IVA)

3.3 Situation awareness

A one-way ANOVA found that participants' transformed SAGAT scores differed significantly across scenarios: $F(3, 69) = 10.44$, $p <$ .001. Thus, transformed SAGAT scores were further converted to Z-scores for each scenario before a two-way repeated ANOVA to understand the influence of speech style and embodiment.

Significant main effects of speech style [$F(1, 23) = 7.52$, $p < .05$, $\eta_p^2 = .25$] and embodiment [$F(1, 23) = 5.05$, $p < .05$, $\eta_p^2 = .18$] were found on participants' standardized SAGAT scores. Participants had a higher situation awareness score when accompanied by informative agents or robot agents than when accompanied by conversational agents or voice-only agents, respectively (Figure 2). There was no interaction effect between speech style and embodiment on participants' situation awareness.

3.4 Eye-tracking measures

Results from two-way repeated-measures ANOVA indicated a significant main effect of embodiment on distraction fixation frequency (Figure 3): $F(1, 23) = 7.71$, $p < .05$, $\eta_p^2 = .25$. However, there was no difference in total distraction fixation duration between voice-only agent conditions (*Mean* = 293.37 sec, *SD* = 45.20 sec) and robot agent conditions (*Mean* = 290.69 sec, *SD* = 61.41 sec). Speech style did not influence the distraction fixation measures. When accompanied by robot agents, participants made more frequent distracting glances (*Mean* = 0.12, *SD* = 0.11) compared to when accompanied by voice-only agents (*Mean* = 0.07, *SD* = 0.06), regardless of the speech style, but the total distraction duration remained similar.

3.5 Takeover performance

Participants did not differentiate in their takeover methods (using a toggle attached to the steering wheel or pressing the brake) across agent conditions: $\chi^2(3) = 2.30$, $p = .51$. Thus, the subsequent takeover performance analysis did not separate these two methods.

3.5.1 Takeover time. Speech style or embodiment did not have a significant main effect on the takeover reaction time. The average takeover time was 1.46 s (*SD* = 0.28s) and 1.40 s (*SD* = 0.21 s) for the informative voice agent and the conversational voice agent, respectively; and was 1.42 s (*SD* = 0.21 s) and 1.42 s (*SD* = 0.22 s) for the informative robot agent and the conversational robot agent, respectively. Participants had a similar takeover reaction time across all conditions.

3.5.2 Takeover quality. For the construction takeover event, there was a significant interaction effect between speech style and embodiment on the maximum speed: $F(1, 23) = 5.32$, $p < .05$, $\eta_p^2 = .19$, and average speed: $F(1, 23) = 5.49$, $p < .05$, $\eta_p^2 = .19$. A simple main effect analysis indicated that when accompanied by voice-only agents, participants had a numerically higher maximum speed ($p = .067$), a significantly higher minimum speed ($p < .05$), and a significantly higher average speed ($p < .05$) when the agent communicated in an informative style compared to a conversational style (Figure 4). Embodiment was found to have a significant main effect on the maximum lateral acceleration, $F(1, 21) = 4.36$, $p < .05$, $\eta_p^2 = .17$. Participants had a larger maximum lateral acceleration under robot conditions (Figure 5a). Speech style had a significant main effect on the standard deviation of lane position (SDLP), $F(1, 22) = 6.32$, $p <$ .05, $\eta_p^2 = .22$. When accompanied by informative agents, participants had a higher SDLP (Figure 5b).

In the non-lane changing takeover events (i.e., jaywalker, fog, and tunnel), speech style or embodiment did not have any significant main effect on the speed-related measures or SDLP. However, speech style showed a significant main effect on the maximum lateral acceleration, $F(1, 22) = 6.32$, $p < .05$, $\eta_p^2 = .22$. Participants had a larger maximum lateral acceleration under informative agent conditions (Figure 6).

4 DISCUSSION

This study investigated the effects of speech style and embodiment of in-vehicle intelligent agents (IVIAs) on drivers' experience and their takeover performance in a conditionally automated driving condition. Results indicate that although robot agents received

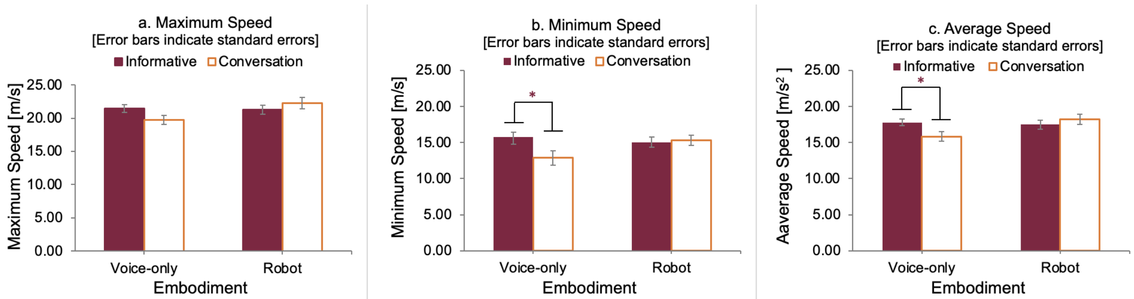

Figure 4: Max/Min/Average speed after takeover for construction (* $p < .05$) [Error bars indicate standard errors]

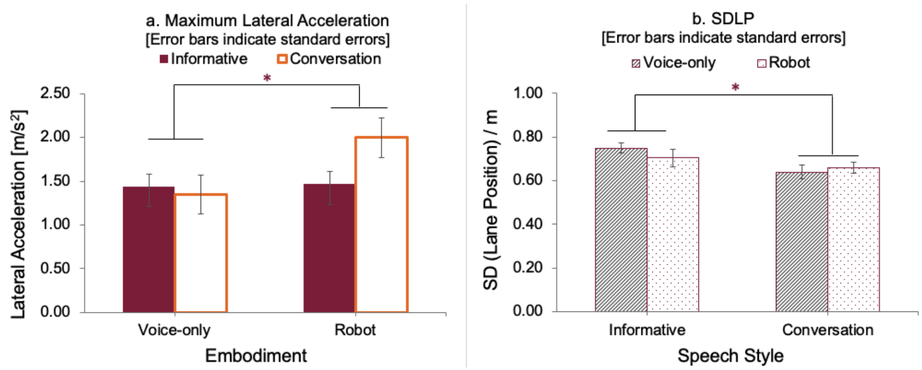

Figure 5: Max lateral acceleration and SDLP after takeover for construction (* $p < .05$) [Error bars indicate standard errors]

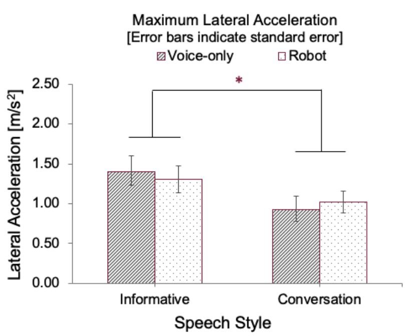

Figure 6: Max lateral acceleration after takeover for other events (* $p < .05$) [Error bars indicate standard errors]

higher ratings in driver-agent interaction questionnaires, they introduced extra distraction, which might cause suboptimal performance after the takeover. On the contrary, conversational agents gained driver preference and demonstrated consistent contribution to safer maneuvers after the takeover, compared to informative agents.

4.1 Speech style and embodiment on driver-agent interaction

While conversational speech style did not outperform informative one, the robot agents were rated more positively on subscales of system response accuracy and trust in automated systems. Although the information presented to drivers was the same between the voice-only and robot conditions, with perfect accuracy, the messages delivered by robot agents were perceived as more accurate and, therefore, met user expectations [17]. Thus, higher trust in the embodied IVIAs was observed in our study as one of the benefits of elevated perceived accuracy that was used to form appropriate trust [3]. A previous study using the same robot found that NAO conditions overall yielded higher trust than voice-only agents, while a social NAO produced the highest trust score [22]. A review paper also pointed out that embodied agents are consistently more trusted [44].

4.2 User preference favored the conversational voice agent

Findings from participants' preference toward four types of agents support our H1. Although participants rated robot agents higher in driver-agent interaction, they preferred voice-only agents the most ($N = 17$). Distraction and discomfort were two dominant perceptions when participants explained their preferences. The autonomous movement of NAO that made it "humanizing" (P4) was perceived as distracting due to its motion and noise. In addition, robot agents' occasional gaze focusing on participants made them uncomfortable. A long time being looked at by a robot counterpart may increase discomfort [42]. Although the differences did not reach a traditional statistically significant level, subjective ratings on the subscale of discomfort (Table 3) were able to support these statements: robot agents were rated numerically higher on discomfort (IRA = 2.30, CRA = 1.99) than voice-only agents (IVA = 1.74, CVA = 1.85). When looking into the ratings on robot agents themselves, it seems that having a conversational speech style can mitigate such discomfort.

In fact, conversational speech style was also preferred the most (N = 17). Both the tone and contents provided by conversational agents were favored. As opposed to informative agents in a perceptually "condescending" style, participants preferred the conversational agents speaking in a "friendly manner", which formed a feeling of being accompanied by "another passenger". In addition, participants preferred the "additional conversational part" that could provide "more context to its (i.e., the vehicle's) decisions". Thus, we could carefully speculate that conversational agents in our study provided higher perceived system transparency by improving drivers' understanding of agents' intentions [2, 37, 55]. However, there was debating about whether the additional information was necessary. Four participants claimed it was unnecessary because drivers were not heavily engaged; thus, they only needed to know the information when human input was required.

Overall, the conversational voice agent (CVA) as the most preferred agent (N= 11) retains the advantages of conversational agents without the disadvantages of robot agents.

4.3 Situation awareness and gaze fixation

We observed an interesting influence of speech style and embodiment on drivers' situation awareness. Although the conversational style was preferred the most and believed to provide additional contents, drivers received lower SAGAT scores when advised by conversational agents compared to when advised by informative agents. Our H2.1 was not supported. Such performance decrement in preferred conditions may result from overreliance [28, 41] and complacency [40]. Drivers were comfortable around conversational agents and satisfied with the information provided. Thus, they might not allocate adequate attention to the scenarios but depend on the agents to deliver information.

Similarly, although robot agents were commented to be distracting and annoying, drivers performed better in SAGAT queries when accompanied by robot agents. A potential explanation is that drivers may be able to prevent task-unrelated distractions. Although a higher distraction fixation frequency was observed under robot agent conditions, the presence of a robot only introduced a minimum level of distraction because the overall distraction duration remained similar. Thus, such distraction was not detrimental enough to compete with the driver's monitoring tasks. In this way, drivers were still able to overcome this interference and took compensatory actions, for example, improving their vigilance to the current situation in this study.

4.4 Conversational speech style produced careful maneuver

The takeover reaction time did not differ across conditions, which was not surprising because the TORs were delivered using the same set of messages. However, the takeover quality for the lane-changing event (i.e., construction) and non-lane-changing events (i.e., jaywalker, fog, and tunnel) was impacted primarily by speech style. At the same time, the embodiment also played a role independently of and dependently on speech style. In general, drivers advised by informative agents exhibit risky and unstable driving behaviors in terms of higher speed across all speed-related measures, a larger standard deviation of lane position in the lane-changing takeover event, and higher maximum lateral acceleration in non-lane-changing takeover events. Only the effects on speed-related measures were moderated by embodiment; the influence of speech style diminished when the agent possessed a physical body. When taking participants' comments into account, the informative speech style was, in general, "annoying" and "irritating", which may create angry drivers who tend to drive faster [19, 48]. Further, the discomfort from the presence of the physical body was so strong that it might weaken or even override the effects of speech style; for instance, participants had a larger maximum lateral acceleration when accompanied by robot agents than voice-only agents.

When taking all takeover quality measures as a big picture, conversational speech style overall yielded greater takeover quality, which supports our H2.2. Although embodied agents were more favored in subjective ratings, speech style was more decisive and powerful in encouraging safer posttakeover driver intervention.

4.5 Implications, limitations, and future work

Findings from the present study are able to provide guidance on designing IVIAs for conditional AVs to deliver road information and issue TORs, which provide evidence on differentiating needs and requirements of IVIAs in vehicles with different levels of automation [51]. The balance between user preference and overreliance needs to be considered to maximize user acceptance while minimizing system misuse. Additional information explaining the system's current action is also critical to building explanatory and transparent IVIAs, which are helpful in forming appropriate mental models [23]. In contrast to embodied IVIAs preferred in full AVs [25, 52], robot agents lost their likability in conditional AVs, where drivers are still required to take actions. Interaction and companionship are no longer necessary in the form of a physical body but can be sufficient in a polite and friendly speaking style without visual distractions. Additionally, delivering information in a natural, equivalent, and easy-going way, such as the conversational speech style, can promote user experience and elicit empathy in drivers, and further lead to cautious driving behaviors. However, if a conversational style is selected to prioritize user acceptance and experience, contents included in the conversation should be carefully drafted to avoid complacency. Future research is needed to manipulate content richness and identify the balance between user acceptance and performance to present IVIAs for conditional AVs.

Even though our study has provided valuable findings leading to promising implications, we acknowledge some limitations worth further exploration. First, 50% of the speech prompts were TORs and were delivered using the same prompts. While we controlled the effect of prompt lengths on takeover reaction times, such a setting may compromise the differentiation between the informative and conversational agents, leading to equivalent momentary reactions to the speech prompt and similar user perceptions of two speech styles in driver-agent interactions. Second, although all driving scenarios were similar in terms of route and environments, elements on the road prior to the SAGAT freeze point differed to some extent. Such differences might introduce the scenario as a covariate when evaluating the SAGAT score. Even though we balanced the match between agent conditions and driving scenarios and standardized scores prior to further analysis, such a variety in the difficulty levels of SAGAT queries might impact user perception in an unforeseeable manner.

As intelligent agents gradually penetrate our daily lives, some of them have been applied to vehicles without any automation, where driver tasks dramatically differ from vehicles with conditional or full automation. As a consequence, IVIAs' responsibilities also shifted. Now that we have studies investigating the social attributes of IVIAs across different levels of automation, future attempts could be made to explore the variability in user perception and preference towards IVIAs across automation levels. In addition, we found a potential emotional reaction towards the IVIAs in the present study. The capabilities of IVIAs to elicit emotional states could be further evaluated and implied to mitigate performance decrements in emotionally impaired drivers [6, 18].

ACKNOWLEDGMENTS

The first author was awarded Doctoral Fellowship from UPS Foundation to support her education and research.

REFERENCES

[1] Adriana Barón and Paul Green. 2006. Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review.

[2] Adella Bhaskara, Michael Skinner, and Shayne Loft. 2020. Agent transparency: A review of current theory and evidence. IEEE Transactions on Human-Machine Systems 50, 3: 215–224. https://doi.org/10.1109/THMS.2020.2965529

[3] Kimberly A. Brink and Henry M. Wellman. 2020. Robot Teachers for Children? Young Children Trust Robots Depending on Their Perceived Accuracy and Agency. Developmental Psychology 56, 7: 1268–1277. https://doi.org/10.1037/dev0000884

[4] Jacob Cohen, Patricia Cohen, Stephen G. West, and Leona S. Aiken. 2013. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Routledge. https://doi.org/10.4324/9780203774441

[5] Henriette Cramer, Vanessa Evers, Nicander Kemper, and Bob Wielinga. 2008. Effects of autonomy, traffic conditions and driver personality traits on attitudes and trust towards In-vehicle agents. Proceedings - 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IAT Workshops 2008 3: 477–482. https://doi.org/10.1109/WIIAT.2008.326

[6] Thomas A. Dingus, Feng Guo, Suzie Lee, Jonathan F. Antin, Miguel Perez, Mindy Buchanan-King, and Jonathan Hankey. 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences of the United States of America 113, 10: 2636–2641. https://doi.org/10.1073/pnas.1513271113

[7] Ebru Dogan, Mohamed-Cherif Rahal, Renaud Deborne, Patricia Delhomme, Andras Kemeny, and Jérôme Perrin. 2017. Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. Transportation Research Part F: Traffic Psychology and Behaviour 46: 205–215. https://doi.org/10.1016/j.trf.2017.01.012

[8] Jiayuan Dong, Emily Lawson, Jack Olsen, and Myounghoon Jeon. 2020. Female voice agents in fully autonomous vehicles are not only more likeable and comfortable, but also more competent. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 64, 1: 1033–1037. https://doi.org/10.1177/1071181320641248

[9] Na Du, Feng Zhou, Dawn Tilbury, Lionel Peter Robert, and X Jessie Yang. 2021. Designing alert systems in takeover transitions: The effects of display information and modality. In Proceedings - 13th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2021, 173–180. https://doi.org/10.1145/3409118.3475155

[10] Mica R. Endsley. 1995. Toward a theory of situation awareness in dynamic systems. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 1: 32–64. https://doi.org/10.1518/001872095779049543

[11] Mica R. Endsley. 2021. Situation Awareness Measurement: How to Measure Situation Awareness in Individuals and Teams. Human Factors and Ergonomics Society.

[12] Mica R. Endsley and Daniel J. Garland. 2000. Situation Awareness Analysis and Measurement. CRC Press, Mahwah, N.J. https://doi.org/10.1201/b12461

[13] M.R. Endsley. Situation awareness global assessment technique (SAGAT). In Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, 789–795. https://doi.org/10.1109/NAECON.1988.195097

[14] Yannick Forster, Frederik Naujoks, and Alexandra Neukum. 2017. Increasing anthropomorphism and trust in automated driving functions by adding speech output. In 2017 IEEE Intelligent Vehicles Symposium (IV), 365–372. https://doi.org/10.1109/IVS.2017.7995746

[15] Thomas M Gable and Bruce N Walker. 2013. Georgia Tech Simulator Sickness Screening Protocol Georgia. Georgia Tech School of Psychology Tech Report GT-PSYC-TR-2013-01: 1–16. Retrieved from http://hdl.handle.net/1853/53375

[16] Michelle Hester, Kevin Lee, and Brian P. Dyre. 2017. "Driver take over": A preliminary exploration of driver trust and performance in autonomous vehicles. Proceedings of the Human Factors and Ergonomics Society 2017-Octob: 1969–1973. https://doi.org/10.1177/1541931213601971

[17] Kate S. Hone and Robert Graham. 2000. Towards a tool for the subjective assessment of speech system interfaces (SASSI). Natural Language Engineering 6, 3–4: 287–291. https://doi.org/10.1017/s1351324900002497

[18] Myounghoon Jeon. 2016. Don't Cry While You're Driving: Sad Driving Is as Bad as Angry Driving. International Journal of Human-Computer Interaction 32, 10: 777–790. https://doi.org/10.1080/10447318.2016.1198524

[19] Myounghoon Jeon, Bruce N. Walker, and Thomas M. Gable. 2015. The effects of social interactions with in-vehicle agents on a driver's anger level, driving performance, situation awareness, and perceived workload. Applied Ergonomics 50: 185–199. https://doi.org/10.1016/j.apergo.2015.03.015

[20] Jiun-Yin Jian, Ann M. Bisantz, and Colin G. Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics 4, 1: 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

[21] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. International Journal on Interactive Design and Manufacturing 9, 4: 269–275. https://doi.org/10.1007/s12008-014-0227-2

[22] Johannes Maria Kraus, Florian Nothdurft, Philipp Hock, David Scholz, Wolfgang Minker, and Martin Baumann. 2016. Human after all: Effects of mere presence and social interaction of a humanoid robot as a co-driver in automated driving. AutomotiveUI 2016 - 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Adjunct Proceedings: 129–134. https://doi.org/10.1145/3004323.3004338

[23] Johannes Maria Kraus, Jessica Sturn, Julian Elias Reiser, and Martin Baumann. 2015. Anthropomorphic agents, transparent automation and driver personality: Towards an integrative multi-level model of determinants for effective driver-vehicle cooperation in highly automated vehicles. In Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15, 8–13. https://doi.org/10.1145/2809730.2809738

[24] David R. Large, Gary Burnett, Kyle Harrington, Leigh Clark, Jacob Luton, Peter Thomas, and Pete Bennett. 2019. "It's small talk, jim, but not as we know it." Engendering trust through human-agent conversation in an autonomous, self-driving car. ACM International Conference Proceeding Series. https://doi.org/10.1145/3342775.3342789

[25] David R. Large, Kyle Harrington, Gary Burnett, Jacob Luton, Peter Thomas, and Pete Bennett. 2019. To please in a pod: Employing an anthropomorphic agent-interlocutor to enhance trust and user experience in an autonomous, self-driving vehicle. Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019: 49–59. https://doi.org/10.1145/3342197.3344545

[26] Dave Lee. 2020. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. Transportation Research Board Annual Meeting 42, 5 June 2020: 1–39.

[27] Jae gil Lee and Kwan Min Lee. 2022. Polite speech strategies and their impact on drivers' trust in autonomous vehicles. Computers in Human Behavior 127, August 2021: 107015. https://doi.org/10.1016/j.chb.2021.107015

[28] John D 1965- Lee, Christopher D Wickens, Yili Liu, and Linda Ng Boyle. 2017. Designing for people: an introduction to human factors engineering. CreateSpace, Charleston, SC SE - xv, 675 pages: illustrations; 26 cm.

[29] John D. Lee and Katrina A. See. 2004. Trust in automation: Designing for appropriate reliance. Human Factors 46, 1: 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

[30] Sanguk Lee, Rabindra Ratan, and Taiwoo Park. 2019. The voice makes the car: Enhancing autonomous vehicle perceptions and adoption intention through voice agent gender and style. Multimodal Technologies and Interaction 3, 1. https://doi.org/10.3390/mti3010020

[31] Seul Chan Lee and Myounghoon Jeon. 2022. A Systematic Review of Functions and Design Features of In-Vehicle Agents. International Journal of Human-Computer Studies: 102864. https://doi.org/10.1016/J.IJHCS.2022.102864

[32] Seul Chan Lee, Sangjin Ko, Harsh Sanghavi, and Myounghoon Jeon. 2019. Autonomous driving with an agent: Speech style and embodiment. Adjunct Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019: 209–214. https://doi.org/10.1145/3349263.3351515
[33] Jannette Maciej and Mark Vollrath. 2009. Comparison of manual vs. speech-based interaction with in-vehicle information systems. Accident Analysis and Prevention 41, 5: 924–930. https://doi.org/10.1016/j.aap.2009.05.007
[34] Kirti Mahajan, David R. Large, Gary Burnett, and Nagendra R. Velaga. 2021. Exploring the benefits of conversing with a digital voice assistant during automated driving: A parametric duration model of takeover time. Transportation Research Part F: Traffic Psychology and Behaviour 80: 104–126. https://doi.org/10.1016/j.trf.2021.03.012
[35] Kirti Mahajan, David R. Large, Gary Burnett, and Nagendra R. Velaga. 2021. Exploring the effectiveness of a digital voice assistant to maintain driver alertness in partially automated vehicles. Traffic Injury Prevention 22, 5: 378–383. https://doi.org/10.1080/15389588.2021.1904138
[36] Anthony D. McDonald, Hananeh Alambeigi, Johan Engström, Gustav Markkula, Tobias Vogelpohl, Jarrett Dunne, and Norbert Yuma. 2019. Toward Computational Simulations of Behavior During Automated Driving Takeovers: A Review of the Empirical and Modeling Literatures. Human Factors 61, 4: 642–688. https://doi.org/10.1177/0018720819829572
[37] Koen van de Merwe, Steven Mallam, and Salman Nazir. 2022. Agent Transparency, Situation Awareness, Mental Workload, and Operator Performance: A Systematic Literature Review. Human Factors: The Journal of the Human Factors and Ergonomics Society: 001872082210778. https://doi.org/10.1177/00187208221077804
[38] Michael A. Nees, Benji Helbein, and Anna Porter. 2016. Speech Auditory Alerts Promote Memory for Alerted Events in a Video-Simulated Self-Driving Car Ride. Human Factors 58, 3: 416–426. https://doi.org/10.1177/0018720816629279
[39] Lin Padgham and Michael Winikoff. 2005. Agents and Multi-Agent Systems. In Developing Intelligent Agent Systems, Michael Wooldridge (ed.). John Wiley & Sons, Ltd, Chichester, UK, 1–6. https://doi.org/10.1002/0470861223.ch1
[40] Raja Parasuraman and Dietrich H. Manzey. 2010. Complacency and bias in human use of automation: An attentional integration. Human Factors 52, 3: 381–410. https://doi.org/10.1177/0018720810376055
[41] Raja Parasuraman and Victor Riley. 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. Human Factors: The Journal of the Human Factors and Ergonomics Society 39, 2: 230–253. https://doi.org/10.1518/001872097778543886
[42] Maria Teresa Parreira and Sarah Gillet. 2022. Design Implications for Effective Robot Gaze Behaviors in Multiparty Interactions. 2017: 976–980.
[43] Ioannis Politis, Stephen Brewster, and Frank Pollick. 2015. Language-based multimodal displays for the handover of control in autonomous cars. c: 3–10. https://doi.org/10.1145/2799250.2799262
[44] Minjin Rheu, Ji Youn Shin, Wei Peng, and Jina Huh-Yoo. 2020. Systematic Review: Trust-Building Factors and Implications for Conversational Agent Design. International Journal of Human-Computer Interaction 00, 00: 1–16. https://doi.org/10.1080/10447318.2020.1807710
[45] Fabienne Roche and Stefan Brandenburg. 2020. Should the Urgency of Visual-Tactile Takeover Requests Match the Criticality of Takeover Situations. IEEE Transactions on Intelligent Vehicles 5, 2: 306–313. https://doi.org/10.1109/TIV.2019.2955906
[46] E. Roesler, D. Manzey, and L. Onnasch. 2021. A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction. Science robotics 6, 58. https://doi.org/10.1126/SCIROBOTICS.ABJ5425
[47] Society of Automotive Engineers. 2018. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems.
[48] Klara Steinhauser, Felix Leist, Kathrin Maier, Vera Michel, Nikolai Pärsch, Philip Rigley, Franz Wurm, and Marco Steinhauser. 2018. Effects of emotions on driving behavior. Transportation Research Part F: Traffic Psychology and Behaviour 59: 150–163. https://doi.org/10.1016/j.trf.2018.08.012
[49] Tobias Vogelpohl, Matthias Kühn, Thomas Hummel, Tina Gehlert, and Mark Vollrath. 2018. Transitioning to manual driving requires additional time after automation deactivation. Transportation Research Part F: Traffic Psychology and Behaviour 55: 464–482. https://doi.org/10.1016/j.trf.2018.03.019
[50] Tobias Vogelpohl, Matthias Kühn, Thomas Hummel, and Mark Vollrath. 2019. Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. Accident Analysis & Prevention 126: 70–84. https://doi.org/10.1016/j.aap.2018.03.013
[51] Manhua Wang, Philipp Hock, Seul Chan Lee, Martin Baumann, and Myounghoon Jeon. 2021. Genie vs. Jarvis: Characteristics and Design Considerations of In-Vehicle Intelligent Agents. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 197–199. https://doi.org/10.1145/3473682.3479720
[52] Manhua Wang, Seul Chan Lee, Harsh Kamalesh Sanghavi, Megan Eskew, Bo Zhou, and Myounghoon Jeon. 2021. In-Vehicle Intelligent Agents in Fully Autonomous Driving: The Effects of Speech Style and Embodiment Together and Separately. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 247–254. https://doi.org/10.1145/3409118.3475142
[53] Yi Wang, Wei Zhang, and Ronggang Zhou. 2022. Speech-based takeover requests in conditionally automated driving: Effects of different voices on the driver takeover performance. Applied Ergonomics 101, August 2021: 103695. https://doi.org/10.1016/j.apergo.2022.103695
[54] Priscilla N.Y. Wong, Duncan P. Brumby, Harsha Vardhan Ramesh Babu, and Kota Kobayashi. 2019. "Watch out!" Semi-autonomous vehicles using assertive voices to grab distracted drivers' attention. Conference on Human Factors in Computing Systems - Proceedings: 5–10. https://doi.org/10.1145/3290607.3312838
[55] Julia L Wright, Jessie YC Chen, Michael J Barnes, and Peter A Hancock. 2017. Agent Reasoning Transparency: The Influence of Information Level on Agent Reasoning Transparency: The Influence of Information Level on Automation-Induced Complacency. June.
[56] Youngjae Yoo, Min young Yang, Seunghoon Lee, Hyungwoo Baek, and Jinwoo Kim. 2022. The effect of the dominance of an in-vehicle agent's voice on driver situation awareness, emotion regulation, and trust: A simulated lab study of manual and automated driving. Transportation Research Part F: Traffic Psychology and Behaviour 86: 33–47. https://doi.org/10.1016/J.TRF.2022.01.009
[57] Maryam Zahabi and David Kaber. 2018. Effect of police mobile computer terminal interface design on officer driving distraction. Applied Ergonomics 67: 26–38. https://doi.org/10.1016/j.apergo.2017.09.006
[58] Yiqi Zhang, Changxu Wu, and Jingyan Wan. 2016. Mathematical Modeling of the Effects of Speech Warning Characteristics on Human Performance and Its Application in Transportation Cyberphysical Systems. IEEE Transactions on Intelligent Transportation Systems 17, 11: 3062–3074. https://doi.org/10.1109/TITS.2016.2539975
[59] Jakub Złotowski, Diane Proudfoot, Kumar Yogeeswaran, and Christoph Bartneck. 2015. Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction. International Journal of Social Robotics 7, 3: 347–360. https://doi.org/10.1007/S12369-014-0267-6/FIGURES/1